PS7

Owen McDevitt

March 2019

Table 1:

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
logwage	1,669	1.625	0.386	0.005	1.362	1.936	2.261
hgc	2,229	13.101	2.524	0	12	15	18
tenure	2,229	5.971	5.507	0.000	1.583	9.333	25.917
age	2,229	39.152	3.062	34	36	42	46

1. The log wages are missing about one quarter of the observations, and I highly doubt they are missing at random. People who do not have wages are very likely to share some common characteristics and self-select out of the workforce. Thus, I think they are MNAR.

2. On the next page is the regression using three different imputation methods (list-wise deletion, mean imputation, and mice). For some reason however, when I run the MICE imputation I get the same results as the mean imputation. I tried many times and kept getting the same results. Regardless though, being married was not a significant factor in list-wise deletion, but it was in the other two. This tells me that many of the people with missing data are probably married, and thus by disregarding them, we lose the real relationship. Unfortunately, the mean imputation and mice got further away from the true value of beta 1, however I attribute that to my inability to run mice correctly.

3. My project for this class is the project with you. Unfortunately, I have been really busy and have not made much progress. I plan on changing that asap.

Table 2:

	Dependent variable:		
	logwage		
	(1)	(2)	(3)
hgc	0.062***	0.049***	0.049***
	(0.005)	(0.004)	(0.004)
collegenot college grad	0.146***	0.160***	0.160***
	(0.035)	(0.026)	(0.026)
tenure	0.023***	0.015***	0.015***
	(0.002)	(0.001)	(0.001)
age	−0.001	−0.001	−0.001
	(0.003)	(0.002)	(0.002)
marriedsingle	−0.024	−0.029**	−0.029**
	(0.018)	(0.014)	(0.014)
Constant	0.639***	0.833***	0.833***
	(0.146)	(0.115)	(0.115)
Observations	1,669	2,229	2,229
R^2	0.195	0.132	0.132
Adjusted R^2	0.192	0.130	0.130
Residual Std. Error	0.346 (df = 1663)	0.311 (df = 2223)	0.311 (df = 2223)
F Statistic	80.508*** (df = 5; 1663)	67.496*** (df = 5; 2223)	67.496*** (df = 5; 2223
Note:			*p<0.1; **p<0.05; ***p<0.0